JOURNAL OF FUNCTIONAL ANALYSIS 95, 414–432 (1991)

Absolute Continuity of Distributions of Solutions of Anticipating Stochastic Differential Equations

TAKAO MASUDA

Kobe University of Commerce, 651-21 Kobe, Japan

Communicated by Paul Malliavin

Received September 1989

We consider a stochastic differential equation with anticipating initial value and drift, and give two types of sufficient conditions for the absolute continuity of the distribution of the solution with respect to the Lebesgue measure, using the Malliavin calculus.

1. INTRODUCTION

Let $w(t)$, $0 \leqslant t \leqslant 1$, be a d-dimensional Brownian motion starting at the origin. Then, the pinned Brownian motion $X_t = w(t) - tw(1)$ can be considered as a solution of the equation

$$dX_t = dw(t) - w(1)\,dt. \tag{1.1}$$

But this is not a stochastic differential equation of usual type because the drift $w(1)$ is not adapted to the filtration $(\mathscr{F}_t)_{0 \leqslant t \leqslant 1}$ where $\mathscr{F}_t$ is a σ-field generated by $\{w(s) : 0 \leqslant s \leqslant t\}$. Recently several authors defined stochastic integrations of anticipating processes. Especially D. Ocone and E. Pardoux considered in [5] an anticipating stochastic differential equation in $\mathbb{R}^d$ of the type

$$X_t = X_0 + \sum_{k=1}^{r} \int_0^t \sigma_k(s, X_s) \circ dw^k(s) + \int_0^t B(s, X_s)\,ds, \tag{1.2}$$

where the σ_k's are non-random vector fields but X_0 and B depend on the full path of Wiener process $w(t) = (w^1(t), \ldots, w^r(t))$, $0 \leqslant t \leqslant 1$, and they proved the well-posedness of (1.2) under a certain condition of smoothness and growth concerning coefficients and initial value (see Theorem 2.9).

In this paper, we study the absolute continuity of the distribution of the

0022-1236/91 $3.00

solution X_t. We define vector fields $\sigma_{k_0 \cdots k_m}$ on $\mathbb{R}^d$ from $\sigma_1, ..., \sigma_r$ and B by induction as

$$\sigma_0 = B$$

$$\sigma_{k_0 \cdots k_m k}(t, x) = [\sigma_k(t), \sigma_{k_0 \cdots k_m}(t)](x), \qquad k = 1, ..., r$$

$$\sigma_{k_0 \cdots k_m 0}(t, x) = [B(t), \sigma_{k_0 \cdots k_m}(t)](x) + \frac{\partial \sigma_{k_0 \cdots k_m}}{\partial t}(t, x),$$

where $[\cdot, \cdot]$ denotes the Lie bracket of vector fields. In the case where the initial value X_0 is a constant x and the drift B is a non-random vector field σ_0, it is known that if

$$\operatorname{rank}\{\sigma_{k_0 \cdots k_m}(0, x) : m \geqslant 0, k_0 \neq 0\} = d, \tag{1.3}$$

then the distribution of X_t is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^d$ for all $t \in (0, 1]$ (cf. [2]). So it is natural to ask in the anticipating case whether the rank condition (1.3) implies the absolute continuity of the distribution of X_t. But it is not the case in general. Indeed, the stochastic differential equation (1.1) satisfies (1.3) obviously but the distribution of X_1 is singular since $X_1 = 0$.

Our results are stated in two theorems. In the first one (Theorem 3.1), we show that the distribution of X_t is absolutely continuous for almost all t if the equation satisfies the rank condition and $B(t, x)$ satisfies some regularity condition. In the second one (Theorem 3.2), we show that the distribution of X_t is absolutely continuous for all $t \in (0, 1]$ if the equation satisfies two conditions;

(i) $B(t, x)$ and X_0 are written as functionals of multiple Wiener–Stratonovich integrals of degree q,

(ii) $\operatorname{rank}\{\sigma_{k_0 \cdots k_m}(0, X_0) : m \geqslant q, k_0 \neq 0\} = d$ a.s.

For the proof to these theorems, the Malliavin calculus [3] is fully used.

2. Stochastic Calculus with Anticipating Processes

In this section, we shall review briefly the notion of stochastic integration introduced in [4] and the result for the well-posedness of the stochastic differential equation (1.2), which was first given in [5].

Let $(\mathbb{W}, \mathbb{H}, P)$ be the r-dimensional standard Wiener space; i.e., $\mathbb{W}$ is the totality of $\mathbb{R}^r$-valued continuous functions $w = (w^1, ..., w^r)$ on the interval $[0, 1]$ such that $w(0) = 0$, $\mathbb{H} = H^r$ where H is the totality of absolutely continuous functions h on $[0, 1]$ such that $h(0) = 0$ and $\int_0^1 |\dot{h}(t)|^2 \, dt < \infty$ where

$\dot{h}$ denotes the derivative of h, and P is the standard Wiener measure on $\mathbb{W}$. We identify H with $L^2([0, 1])$ by identifying h in H with $\dot{h}$ in $L^2([0, 1])$ and we consider H as a Hilbert space. We also identify the n-tensor product $H^{\otimes n}$ with $L^2([0, 1]^n)$. Through this paper, all random variables and stochastic processes are assumed to be defined on $\mathbb{W}$.

We shall define the Sobolev spaces $\mathbb{D}^{n,p}(E)$ of Wiener functionals and $\mathbb{H}$-derivation $D=(D_1, ..., D_r)$ as follows (cf. [6, 7]). We call a random variable F smooth if F has the form

$$F(w)=f(w^{k_1}(t_1), ..., w^{k_n}(t_n)),$$

where $n \geqslant 0$, $f \in C_b^\infty(\mathbb{R}^n)$, $1 \leqslant k_1, ..., k_n \leqslant r$, $0 \leqslant t_1 < \cdots < t_n \leqslant 1$. Then, we define an $\mathbb{H}^{\otimes n}$-valued random variable $D^n F=(D^n_{l_1 \cdots l_n} F)_{1 \leqslant l_1, ..., l_n \leqslant r}$ as follows: if $(\dot{D}^n_{l_1 \cdots l_n} F)(w)$ denotes the element of $L^2([0, 1]^n)$ which is identified with $(D^n_{l_1 \cdots l_n} F)(w) \in H^{\otimes n}$, then

$$\begin{aligned}((\dot{D}^n_{l_1 \cdots l_n} F)(w))(\lambda_1, ..., \lambda_n) \\ = \sum \frac{\partial^n f}{\partial x_{i_1} \cdots \partial x_{i_n}} ((w^{k_i}(t_i))_i) \chi_{[0, t_{i_1}]}(\lambda_1) \cdots \chi_{[0, t_{i_n}]}(\lambda_n),\end{aligned}$$

where we take the sum over $i_1, ..., i_n$ such that $l_{i_1}=k_1, ..., l_{i_n}=k_n$. $D^n F$ is characterized by

$$\begin{aligned}\frac{\partial^n}{\partial \varepsilon_1 \cdots \partial \varepsilon_n} F(w+\varepsilon_1 h_1 + \cdots + \varepsilon_n h_n)|_{\varepsilon_1 = \cdots = \varepsilon_n = 0} \\ = (D^n F(w), h_1 \otimes \cdots \otimes h_n)_{\mathbb{H}^{\otimes n}} \qquad \text{for any} \quad h_1, ..., h_n \in \mathbb{H}.\end{aligned}$$

Then the operator $D^n: S \subset L^2(\mathbb{W}) \to L^2(\mathbb{W}; \mathbb{H}^{\otimes n})$ is closable (cf. [6]), where S denotes the totality of smooth random variables. For $F \in S$, set

$$\|F\|_{\mathbb{D}^{n,p}} = \left(\mathbb{E}\left[|F|^p + \sum_{m=1}^{n} \|D^m F\|^p_{\mathbb{H}^{\otimes m}} \right] \right)^{1/p} \qquad n=0, 1, 2, ..., p \geqslant 2$$

and let $\mathbb{D}^{n,p}$ be the $\| \ \|_{\mathbb{D}^{n,p}}$-completion of S. For a separable Hilbert space E, we set $\mathbb{D}^{n,p}(E)=\mathbb{D}^{n,p} \otimes E \subset L^2(\mathbb{W}, E)$. Then, naturally the operators

$$D^m: \mathbb{D}^{n,p}(E) \to \mathbb{D}^{n-m,p}(\mathbb{H}^{\otimes m} \otimes E)$$

and

$$\dot{D}^m: \mathbb{D}^{n,p}(E) \to \mathbb{D}^{n-m,p}(L^2([0, 1]^m; E^{r^m}))$$

are defined. Moreover they become bounded operators (cf. [7]).

Next, we give a definition of stochastic integration of anticipating processes following Nualart and Pardoux [4] and prepare some basic properties.

DEFINITION 2.1. Let $n=1, 2, ..., p \geqslant 2$, and E be a separable Hilbert space. We say that an E-valued stochastic process $\{u_t\}_{0\leqslant t\leqslant 1}$ belongs to $\tilde{\mathbb{L}}^{n,p}(E)$ if

(i) $u_t(w)$ is measurable with respect to $(t, w) \in [0, 1] \times \mathbb{W}$,

(ii) $u_t \in \mathbb{D}^{n,p}(E)$ for all $t \in [0, 1]$,

(iii) $\int_0^1 \|u_t\|_{\mathbb{D}^{n,p}}^p \, dt < +\infty$,

(iv) there exists a version of $(\dot{D}^m u_t(w))(\lambda_1, ..., \lambda_m)$ for $m=1, ..., n$ which is measurable with respect to $(t, \lambda_1, ..., \lambda_m, w)$ and satisfies that

(a) for almost every $(\lambda_1, ..., \lambda_m) \in [0, 1]^m$, $(\dot{D}^m u_t)(\lambda_1, ..., \lambda_m)$ converges in $L^p(\mathbb{W}, E^{r^m})$ as $t \downarrow \lambda_1$ and $t \uparrow \lambda_1$,

(b) $\int_0^1 \cdots \int_0^1 \sup_{0\leqslant t\leqslant 1} \mathbb{E}[\|(\dot{D}^m u_t)(\lambda_1, ..., \lambda_m)\|_{E^{r^m}}^p] \, d\lambda_1 \cdots d\lambda_m < \infty$.

The class $\tilde{\mathbb{L}}^{n,p}$ is slightly different from $\mathbb{L}_c^{n,p}$ which was introduced in [4]. But the arguments in [4] still work when we replace $\mathbb{L}_C^{n,p}$ in [4] by $\tilde{\mathbb{L}}^{n,p}$ and we can see the following lemma.

LEMMA 2.2. *Let* $0 \leqslant a < b \leqslant 1$ *and* $u = \{u(t)\}_{0\leqslant t\leqslant 1} \in \tilde{\mathbb{L}}^{1,p}(E)$. *For a partition* $\Delta : 0 = t_0 < t_1 < \cdots < t_l = 1$ *and* $k = 1, ..., r$, *we set*

$$I_{k,(a,b]}^{\Delta}(u) = \sum_{m=1}^{j} \left(\frac{1}{s_m - s_{m-1}} \int_{s_{m-1}}^{s_m} u(t) \, dt \right) (w^k(s_m) - w^k(s_{m-1})), \quad (2.1)$$

where $a = s_0 < s_1 < \cdots < s_j = b$ *and* $\{s_1, ..., s_{j-1}\} = \{t_0, ..., t_l\} \cap (a, b)$. *Then,* $I_{k,(a,b]}^{\Delta}(u)$ *converges in* $L^p(\mathbb{W}; E)$ *as* $|\Delta| \to 0$.

DEFINITION 2.3. We call the limit in Lemma 2.2 the stochastic integral (of Stratonovich type) of u and write $\int_a^b u(t) \circ dw^k(t)$.

The classes $\mathbb{D}^{n,p}$ and $\tilde{\mathbb{L}}^{n,p}$ are too small to consider anticipating stochastic differential equations, so we shall introduce the larger classes $\mathbb{D}_{loc}^{n,p}$ and $\tilde{\mathbb{L}}_{loc}^{n,p}$. We know that if F and G belong to $\mathbb{D}^{1,2}(E)$, then $P(F=G$ and $DF \neq DG) = 0$ (cf. [4]). So we can define the following notion.

DEFINITION 2.4. We say that an E-valued random variable F belongs to $\mathbb{D}_{loc}^{n,p}(E)$ when there exist a sequence $\{F_j\}_{j=1}^{\infty}$ in $\mathbb{D}^{n,p}(E)$ and an increasing sequence $\{\mathbb{W}_j\}_{j=1}^{\infty}$ of measurable subsets of $\mathbb{W}$ such that $P(\bigcup_{j=1}^{\infty} \mathbb{W}_j) = 1$

and $F=F_j$ a.s. on $\mathbb{W}_j$. Then, we define the $\mathbb{H}$-derivative $DF\in\mathbb{D}^{n-1}_{loc}(\mathbb{H}\otimes E)$ by $DF=DF_j$ on $\mathbb{W}_j$.

DEFINITION 2.5. We say that an E-valued stochastic process $u=\{u(t)\}_{0\leqslant t\leqslant 1}$ belongs to $\tilde{\mathbb{L}}^{n,p}_{loc}(E)$ when there exist a sequence $\{u_j\}_{j=1}^{\infty}$ in $\tilde{\mathbb{L}}^{n,p}(E)$ and an increasing sequence $\{\mathbb{W}_j\}_{j=1}^{\infty}$ of measurable subsets of $\mathbb{W}$ such that $P(\bigcup_{j=1}^{\infty}\mathbb{W}_j)=1$ and $u(t)=u_j(t)$ for all $t\in[0,1]$ a.s. on $\mathbb{W}_j$. Then we define $\int_a^b u(t)\circ dw^k(t)$ by $\int_a^b u(t)\circ dw^k(t)=\int_a^b u_j(t)\circ dw^k(t)$ on $\mathbb{W}_j$.

For $u\in\tilde{\mathbb{L}}^{n,p}_{loc}(E)$, it is clear that $\int_a^b u(t)\circ dw^k(t)$ is the limit in probability of $I^{\Delta}_{k,(a,b]}(u)$ when $|\Delta|\to 0$ where $I^{\Delta}_{k,(a,b]}(u)$ is defined by (2.1).

Now we state a criterion for the absolute continuity of the distribution of $\mathbb{R}^d$-valued random variables.

LEMMA 2.6. *Let $F=(F^1,\ldots,F^d)\in\mathbb{D}^{2,2}_{loc}(\mathbb{R}^d)$. We put $\tilde{\mathbb{W}}=\{\det((DF^i,DF^j)_{\mathbb{H}})>0\}$ and let $\tilde{P}$ be the restriction of P on $\tilde{\mathbb{W}}$. Then the measure $\tilde{P}\circ F^{-1}$ is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^d$.*

Proof. It suffices to show the assertion when $F\in\mathbb{D}^{2,2}(\mathbb{R}^d)$. But it can be obtained exactly in the same way as the proof of Theorem 3.1 in [6] so we omit the detail. ∎

The following lemma is a modification of Theorem 7.9 in [4] and can be proved similarly.

LEMMA 2.7 (Itô's Formula). *Let E be a separable Hilbert space and f be a C^2-function on E in the sense of Fréchet such that its second derivative is bounded on any bounded subsets of E. Let E-valued stochastic processes $\{X_t\}_{0\leqslant t\leqslant 1}$, $\{A(t)\}_{0\leqslant t\leqslant 1}$, and $\{B_k(t)\}_{0\leqslant t\leqslant 1}$, $k=1,\ldots,r$, satisfy that*

(i) $X_t=X_0+\int_0^t A(s)\,ds+\sum_{k=1}^r\int_0^t B_k(s)\circ dw^k(s)$ *a.s.* $0\leqslant t\leqslant 1$.

(ii) $\{X_t\}_{0\leqslant t\leqslant 1}$ belongs to $\tilde{\mathbb{L}}^{1,4}_{loc}(E)$ *and is continuous a.s.*

(iii) $\{A(t)\}_{0\leqslant t\leqslant 1}$ *is measurable on* $[0,1]\times\mathbb{W}$ *and all its paths belong to* $L^2([0,1];E)$.

(iv) $\{B_k(t)\}_{0\leqslant t\leqslant 1}$ *belongs to* $\tilde{\mathbb{L}}^{1,4}_{loc}(E)$ *for* $k=1,\ldots,r$.

Then, $\{\langle\partial f(X_t),B_k(t)\rangle_E\}_{0\leqslant t\leqslant 1}\in\tilde{\mathbb{L}}^{1,4}_{loc}(\mathbb{R})$ *and*

$$f(X_t)=f(X_0)+\int_0^t\langle\partial f(X_s),A(t)\rangle_E\,ds$$

$$+\sum_{k=1}^r\int_0^t\langle\partial f(X_s),B_k(s)\rangle_E\circ dw^k(s)\qquad a.s.$$

for all $t\in[0,1]$ *where* $\partial f\colon E\to E$ *denotes the Fréchet derivative of* f.

Next, we shall compute the $\mathbb{H}$-derivative of a stochastic integral.

LEMMA 2.8. *We assume that* $u=\{u(t)\}_{0\leqslant t\leqslant 1}\in\widetilde{\mathbb{L}}^{2,2}_{loc}(E)$. *Then,* $\int_a^b u(t)\circ dw^k(t)\in\mathbb{D}^{1,2}_{loc}(E)$ *for* $0\leqslant a<b\leqslant 1$ *and* $k=1, ..., r$ *and*

$$\dot{D}_l\left(\int_a^b u(t)\circ dw^k(t)\right)=\int_a^b \dot{D}_l u(t)\circ dw^k(t)+\delta_{kl}\chi_{(a,b]}u \qquad a.s. \tag{2.2}$$

holds in $L^2([0,1], E)$.

Proof. It suffices to show the assertion when $u\in\widetilde{\mathbb{L}}^{2,2}(E)$. We define $I^{\Delta}_{k,(a,b]}(u)$ by (2.1). Then, $I^{\Delta}_{k,(a,b]}(u)\to\int_a^b u(t)\circ dw^k(t)$ in $L^2(\mathbb{W}; E)$ as mentioned in Lemma 2.2 and

$$\begin{aligned}\dot{D}_l I^{\Delta}_{k,(a,b]}(u)&=\sum_{m=1}^{j}\left(\frac{1}{s_m-s_{m-1}}\int_{s_{m-1}}^{s_m}\dot{D}_l u(t)\,dt\right)(w^k(s_m)-w^k(s_{m-1}))\\&\quad+\delta_{kl}\sum_{m=1}^{j}\left(\frac{1}{s_m-s_{m-1}}\int_{s_{m-1}}^{s_m}u(t)\,dt\right)\chi_{[s_{m-1},s_m]}\\&\to\int_a^b \dot{D}_l u(t)\circ dw^k(t)+\delta_{kl}\chi_{[a,b]}u\end{aligned}$$

in $L^2(\mathbb{W}; L^2([0,1]; E))$ as $|\Delta|\to 0$. Therefore, we can conclude the assertion of the lemma from the closedness of the operator D_l. ∎

In the remainder of this section, we state the well-posedness of the stochastic differential equation (1.2) which is a modification of the result in [5] and can be obtained following the idea in [5]. Let B be represented as $B=\tilde{\sigma}_0+\widetilde{B}$ where $\tilde{\sigma}_0$ is a non-random vector field, and we introduce the following assumptions (A.1)–(A.7):

(A.1) $\sigma_1, ..., \sigma_r, \tilde{\sigma}_0\in C^\infty([0,1]\times\mathbb{R}^d; \mathbb{R}^d)$. Moreover, $\sigma_1, ..., \sigma_r$ and $\tilde{\sigma}_0+(1/2)\sum_{k=1}^r\partial\sigma_k\cdot\sigma_k$ have bounded derivatives of any order where $\partial\sigma_k(t,x)$ denotes the Jacobi matrix $((\partial\sigma^i_k/\partial x^j)(t,x))_{i,j=1,...,d}$.

(A.2) $\widetilde{B}(\cdot,\cdot,w)\in C^\infty([0,1]\times\mathbb{R}^d; \mathbb{R}^d)$ for all $w\in\mathbb{W}$.

(A.3) There is a positive function ε on $\mathbb{W}$ such that

$$\sup_{0\leqslant t\leqslant 1, x\in\mathbb{R}^d}\frac{|\widetilde{B}(t,x,w)|}{(1+|x|)^{1-\varepsilon(w)}}<\infty \qquad \text{for all} \quad w\in\mathbb{W}.$$

(A.4) For any $K>0$ and $n=0, 1, 2, ...$, it holds that

$$\sup\{|\partial^n\widetilde{B}(t,x,w)| : 0\leqslant t\leqslant 1, |x|\leqslant K, w\in\mathbb{W}\}<\infty,$$

where ∂ denotes the derivation in x.

(A.5) $\tilde{B}(t, x, \cdot) \in \bigcap_p \mathbb{D}^{2,p}(\mathbb{R}^d)$ for all t and x.

(A.6) For $m = 1, 2$, there exists a version of $(\dot{D}^m \tilde{B}(t, x, w))(\lambda)$, $t \in [0, 1]$, $x \in \mathbb{R}^d$, $w \in \mathbb{W}$, $\lambda \in [0, 1]^m$ such that $(\dot{D}^m \tilde{B}(t, \cdot, w))(\lambda) \in C^\infty(\mathbb{R}^d; \mathbb{R}^d)$ and

$$\int_{[0,1]^m} E[\sup_{0 \leqslant t \leqslant 1, |x| \leqslant K} |\partial^n \dot{D}^m \tilde{B}(t, x, \lambda)|^p] \, d\lambda < \infty$$

for any $K > 0$, $n = 1, 2, \ldots$.

(A.7) $X_0 \in \bigcap_p \mathbb{D}^{2,p}(\mathbb{R}^d)$.

Let $\varphi_t(x)$, $0 \leqslant t \leqslant 1$, $x \in \mathbb{R}^d$, be the stochastic flow of diffeomorphisms defined by the non-anticipating stochastic differential equation

$$\varphi_t(x) = x + \sum_{k=1}^{r} \int_0^t \sigma_k(s, \varphi_s(x)) \circ dw^k(s) + \int_0^t \tilde{\sigma}_0(s, \varphi_s(x)) \, ds$$

and consider a (random) ordinary differential equation

$$\frac{dZ_t}{dt} = (\partial \varphi_t(Z_t))^{-1} \tilde{B}(t, \varphi_t(Z_t)), \qquad Z_0 = X_0. \tag{2.3}$$

THEOREM 2.9 [6]. *Under the assumptions* (A.1)–(A.7), *there exists a global solution* $\{Z_t\}_{0 \leqslant t \leqslant 1}$ *to* (2.3) *and if we put* $X_t = \varphi_t(Z_t)$, *then* $\{X_t\}_{0 \leqslant t \leqslant 1}$ *is a unique solution to* (1.2) *which has continuous paths and belongs to* $\bigcap_p \tilde{\mathbb{L}}^{2,p}_{loc}(\mathbb{R}^d)$.

3. ABSOLUTE CONTINUITY

We shall give sufficient conditions for the absolute continuity of the distribution of the solution to (1.2) in this section (Theorems 3.1 and 3.2). We assume (A.1)–(A.7) so that there exists a unique solution $\{X_t\}_{0 \leqslant t \leqslant 1}$ to (1.2) which has continuous paths and belongs to $\bigcap_p \tilde{\mathbb{L}}^{2,p}_{loc}(\mathbb{R}^d)$ by Theorems 2.9.

THEOREM 3.1. *Let* $\sigma_{k_0 \cdots k_m} \in C^\infty([0, 1] \times \mathbb{R}^d; \mathbb{R}^d)$, $m \geqslant 0$, $0 \leqslant k_0, \ldots, k_m \leqslant r$. *be the vector fields defined in Section* 1. *We assume that there exists a null set* N *such that*

$$\dim \operatorname{span}\{(DB^i(t, x))(w) \colon 0 \leqslant t \leqslant 1, x \in \mathbb{R}^d, 1 \leqslant i \leqslant d\} < \infty$$

$$\textit{for all} \quad w \in \mathbb{W} - N$$

and

$$\operatorname{rank}\{\sigma_{k_0\cdots k_m}(t, x, w): m \geqslant 0, 1 \leqslant k_0 \leqslant r, 0 \leqslant k_1, ..., k_m \leqslant r\} = d$$

for all $x \in \mathbb{R}^d$, $w \in \mathbb{W} - N$ *and t's belonging to a dense subset of* $[0, 1]$. *Then the distribution of* X_t *is absolutely continuous with respect to the Lebesgue measure on* $\mathbb{R}^d$ *for a.e. t in* $[0, 1]$. *Moreover, if*

$$\sup_{w \in W} \dim \operatorname{span}\{(DB^i(t, x))(w) : 0 \leqslant t \leqslant 1, x \in \mathbb{R}^d, 1 \leqslant i \leqslant d\} < \infty,$$

then the number of t's such that the distribution of X_t *is not absolutely continuous with respect to the Lebesgue measure on* $\mathbb{R}^d$ *is at most countable.*

In order to state another theorem, we introduce some classes of H-valued random variables. Let us fix a nonnegative integer q. Let e be a positive integer and $\xi \in \mathbb{R}^e$. Then consider a stochastic differential equation

$$\xi_t = \xi + \sum_{k=1}^{r} \int_0^t V_k(s, \xi_s) \circ dw^k(s) + \int_0^t V_0(s, \xi_s)\, ds \tag{3.1}$$

in $\mathbb{R}^e$ where $V_0, V_1, ..., V_r \in C^\infty([0, 1] \times \mathbb{R}^e; \mathbb{R}^e)$ satisfy two conditions:

(i) $V_1, ..., V_r$ and $V_0 + (1/2) \sum_{k=1}^r \partial V_k \cdot V_k$ have bounded derivatives of any order,

(ii) $V_{k_0\cdots k_m}(0, \xi) = 0$ whenever $m \geqslant q$ and $k_0 \neq 0$, where we define vector fields $V_{k_0\cdots k_m}$ from $V_0, V_1, ..., V_r$ in the way by which we defined $\sigma_{k_0\cdots k_m}$ from $\sigma_0 = B$, $\sigma_1, ..., \sigma_r$.

Let $\{\xi_t = (\xi_t^1, ..., \xi_t^e)\}_{0 \leqslant t \leqslant 1}$ be the solution to (3.1). Then, for each $t \in [0, 1]$ and $k = 1, ..., r$, $G \equiv D_k \xi_t^1$ belongs to $\bigcap_p \mathbb{D}^{1,p}(H)$. We denote by Γ_q the totality of G constructed in this way.

THEOREM 3.2. *We assume that there exists a nonnegative integer q satisfying that*

$$\operatorname{rank}\{\sigma_{k_0\cdots k_m}(0, X_0) : m \geqslant q, 1 \leqslant k_0 \leqslant r,$$
$$0 \leqslant k_1, ..., k_m \leqslant r\} = d \qquad a.s.$$

and there exist finitely many $G^1, ..., G^N \in \Gamma_q$ *such that*

$$D_l X_0^i(w), \qquad (D_l B^i(t, x))(w) \in \operatorname{span}\{G^1(w), ..., G^N(w)\}$$

for all $l = 1, ..., r$, $i = 1, ..., d$, $w \in W$, $t \in [0, 1]$, $x \in \mathbb{R}^d$. *Then the distribututation of* X_t *is absolutely continuous with respect to the Lebesgue measure on* $\mathbb{R}^d$ *for all* $t \in (0, 1]$.

Remark. If X_0 and $B(t, x)$ are functionals of multiple Wiener–Stratonovich integrals $w^{k_1 \cdots k_m}(t_j)$, $m \leqslant q$, $1 \leqslant k_1, ..., k_m \leqslant r$, $1 \leqslant j \leqslant n$, for some $t_1, ..., t_n \in (0, 1]$ where

$$w^{k_1 \cdots k_m}(t) = \int_0^t \int_0^{t_{m-1}} \cdots \int_0^{t_2} w^{k_1}(t_1) \circ dw^{k_2}(t_1) \circ \cdots \circ dw^{k_m}(t_{m-1}),$$

then $D_t X_0^i(w)$ and $(D_t B^i(t, x))(w)$ belong to the linear subspace spanned by the $D_t w^{k_1 \cdots k_m}(t_j)$'s. Then, since the $\mathbb{R}^{r + \cdots + r^q}$-valued process $\{(w^{k_1 \cdots k_m}(t))_{m \leqslant q, 1 \leqslant k_1, ..., k_m \leqslant r}\}_{0 \leqslant t \leqslant 1}$ is written as a solution of the stochastic differential equation of type (3.1), the second condition of Theorem 3.2 is satisfied.

For the proof of absolute continuity of the distribution of X_t, by Lemma 2.6 it is sufficient to show that almost surely the Malliavin covariance $\det((DX_t^i, DX_t^j)_{\mathbb{H}})$ is positive; i.e., $DX_t^1, ..., DX_t^d$ are linearly independent. To do this, we first consider the representation for DX_t, which will be shown through Lemmas 3.3 and 3.4.

LEMMA 3.3. *We consider an anticipating stochastic differential equation*

$$Y_t = I + \sum_{k=1}^{r} \int_0^t \partial\sigma_k(s, X_s) Y_s \circ dw^k(s) + \int_0^t \partial B(s, X_s) Y_s \, ds \tag{3.2}$$

on $\mathbb{R}^d \otimes \mathbb{R}^d$ (the totality of $d \times d$-matrices) where I denotes the unit matrix. Then, (3.2) *has a solution $\{Y_t\}_{0 \leqslant t \leqslant 1}$ which has continuous paths and belongs to $\bigcap_p \tilde{\mathbb{L}}_{loc}^{1,p}(\mathbb{R}^d \otimes \mathbb{R}^d)$. Moreover,* $\det Y_t \neq 0$ *for all t and $\{Y_t^{-1}\}_{0 \leqslant t \leqslant 1}$ satisfies the stochastic differential equation*

$$Y_t^{-1} = I - \sum_{k=1}^{r} \int_0^t Y_s^{-1} \partial\sigma_k(s, X_s) \circ dw^k(s) - \int_0^t Y_0^{-1} \partial B(s, X_s) \, ds. \tag{3.3}$$

Proof. We can construct a solution $\{(X_t, Y_t)\}_{0 \leqslant t \leqslant 1}$ to (1.2) and (3.2) by the method in Theorem 2.9, using the stochastic flow $\{(\varphi_t(x), \partial\varphi_t(x) y) : 0 \leqslant t \leqslant 1, x \in \mathbb{R}^d, y \in \mathbb{R}^d \otimes \mathbb{R}^d\}$ on $\mathbb{R}^d \times (\mathbb{R}^d \otimes \mathbb{R}^d)$. We get a solution to (3.3) similarly and we know it is nothing but $\{Y_t^{-1}\}_{0 \leqslant t \leqslant 1}$ by Ito's formula. ∎

Now we give a representation for DX_t.

LEMMA 3.4. *Let $\{Y_t\}_{0 \leqslant t \leqslant 1}$ be the solution to* (3.2) *introduced in Lemma* 3.3. *Then*

$$DX_t = F_t + G_t \qquad a.s., \tag{3.4}$$

where $F_t = (F_{1,t}, ..., F_{r,t})$ *and* $G_t = (G_{1,t}, ..., G_{r,t})$ *are the elements of* $\bigcap_p \mathbb{D}^{1,p}_{loc}(\mathbb{H} \otimes \mathbb{R}^d)$ *defined by*

$$\dot{F}_{l,t}(\lambda) = Y_t Y_\lambda^{-1} \sigma_l(\lambda, X_\lambda) \chi_{[0,t]}(\lambda) \tag{3.5}$$

$$\dot{G}_{l,t}(\lambda) = Y_t \left\{ (\dot{D}_l X_0)(\lambda) + \int_0^t Y_s^{-1} (\dot{D}_l B(s, x))(\lambda) |_{x = X_s} ds \right\}. \tag{3.6}$$

If X_0 is a constant x and B is a non-random vector field σ_0, it is obvious that $G_t = 0$. In such case, J. M. Bismut showed the linear independency of $DX_t^1, ..., DX_t^d$ [1]. We will show a similar fact by using the quadratic variation of anticipating processes.

Proof of Proposition 3.4. Take the $\mathbb{H}$-derivative of each side of (1.2), then

$$\begin{aligned} \dot{D}_l X_t = \dot{D}_l X_0 &+ \sum_{k=1}^r \int_0^t \partial\sigma_k(s, X_s) \dot{D}_l X_s \circ dw^k(s) \\ &+ \sigma_l(\cdot, X_\cdot) \chi_{[0,t]}(\cdot) + \int_0^t \partial B(s, X_s) \dot{D}_l X_s \, ds \\ &+ \int_0^t \dot{D}_l B(s, x) |_{x = X_s} ds \qquad \text{a.s.} \end{aligned}$$

by Lemma 2.8. So it follows that

$$\begin{aligned} (D_l X_t, h)_H = (D_l X_0, h)_H &+ \sum_{k=1}^r \int_0^t \partial\sigma_k(s, X_s)(D_l X_s, h)_H \circ dw^k(s) \\ &+ \int_0^t \sigma_l(\lambda, X_\lambda) \dot{h}(\lambda) \, d\lambda + \int_0^t \partial B(s, X_s)(D_l X_s, h)_H \, ds \\ &+ \int_0^t (D_l B(s, x), h)_H |_{x = X_s} ds \qquad \text{a.s.} \end{aligned}$$

for any $h \in H$. Noting (3.3), we have by Itô's formula that

$$\begin{aligned} Y_t^{-1}(D_l X_t, h)_H &= (D_l X_0, h)_H + \int_0^t Y_s^{-1} \sigma_l(s, X_s) \dot{h}(s) \, ds \\ &\quad + \int_0^t Y_s^{-1} (D_l B(s, x), h)_H \, ds \\ &= Y_t^{-1} \{ (F_{l,s}, h)_H + (G_{l,s}, h)_H \} \qquad \text{a.s.} \end{aligned}$$

Since H is separable, we have

$$Y_t^{-1} D_l X_t = Y_t^{-1} \{F_{l,t} + G_{l,t}\} \qquad \text{a.s.}$$

and it implies (3.4). ▮

Next, we shall study the linear independency of $\{F_t^i\}_{i,t}$. For this purpose, we set

$$Q_{k_0 \cdots k_m}(\lambda) = Y_\lambda^{-1} \sigma_{k_0 \cdots k_m}(\lambda, X_\lambda)$$

so that $\dot{F}_{l,t}(\lambda) = Y_t Q_l(\lambda) \chi_{[0,t]}(\lambda)$. Noting (1.2) and (3.3), we have

$$Q_{k_0 \cdots k_m}(\lambda) = \sigma_{k_0 \cdots k_m}(0, X_0) + \sum_{k=1}^{r} \int_0^\lambda Q_{k_0 \cdots k_m k}(t) \circ dw^k(t) + \int_0^\lambda Q_{k_0 \cdots k_m 0}(t)\, dt$$

by Itô's formula. Then,

$$\sum_{k=1}^{r} I^{\Delta}_{k,(a,b]}(Q^i_{k_0 \cdots k_m k}) \to \sum_{k=1}^{r} \int_a^b Q^i_{k_0 \cdots k_m k}(t) \circ dw^k(t)$$

$$= Q^i_{k_0 \cdots k_m}(b) - Q^i_{k_0 \cdots k_m}(a) - \int_a^b Q^i_{k_0 \cdots k_m 0}(t)\, dt$$

and

$$[Q^i_{k_0 \cdots k_m}, Q^j_{k_0 \cdots k_m}]^{\Delta}_{(a,b]} \to \sum_{k=1}^{r} \int_a^b Q^i_{k_0 \cdots k_m k}(t) Q^j_{k_0 \cdots k_m k}(t)\, dt$$

in probability as $|\Delta| \to 0$. (See Appendix 1 for latter convergence.) Therefore we can choose a sequence $\{\Delta_n\}_{n=1}^{\infty}$ of partitions of $[0, 1]$ such that

$$\lim_{n \to \infty} \sum_{k=1}^{r} I^{\Delta_n}_{k,(a,b]}(Q^i_{k_0 \cdots k_m k})$$

$$= Q^i_{k_0 \cdots k_m}(b) - Q^i_{k_0 \cdots k_m}(a) - \int_a^b Q_{k_0 \cdots k_m 0}(t)\, dt \qquad (3.7)$$

$$\lim_{n \to \infty} [Q^i_{k_0 \cdots k_m}, Q^j_{k_0 \cdots k_m}]^{\Delta_n}_{(a,b]}$$

$$= \sum_{k=1}^{r} \int_a^b Q^i_{k_0 \cdots k_m k}(t) Q^j_{k_0 \cdots k_m k}(t)\, dt \qquad (3.8)$$

for all rational numbers a and b $(0 \leqslant a < b \leqslant 1)$ where the exceptional set is not merely a null set but empty (as we choose and fix a good version for the $Q^i_{k_0\cdots k_m}$'s).

We show a pathwise property of $Q_{k_0\cdots k_m}(t)$'s which follows from the abovementioned pathwise properties (3.7) and (3.8).

LEMMA 3.5. *Let* $w\in\mathbb{W}$, $v=(v_1, ..., v_d)\in\mathbb{R}^d$, $0\leqslant a<b\leqslant 1$, *and*

$$vQ_l(\lambda, w)=\sum_i v_i Q^i_l(\lambda, w)=0, \qquad a\leqslant\lambda\leqslant b$$

holds. Then it holds that

$$vQ_{lk_1\cdots k_m}(\lambda, w)=0, \qquad a\leqslant\lambda\leqslant b, \quad m\geqslant 1, \quad 0\leqslant k_1, ..., k_m\leqslant r.$$

Proof. For any rational numbers α and β such that $a\leqslant\alpha<\beta\leqslant b$, we have

$$\begin{aligned}
&\sum_{k=1}^{r}\int_\alpha^\beta |vQ_{lk}(\lambda, w)|^2\,d\lambda\\
&\quad=\sum_{k,i,j} v_iv_j\int_\alpha^\beta Q^i_{lk}(\lambda, w)Q^j_{lk}(\lambda, w)\,d\lambda\\
&\quad=\lim_{n\to\infty} v_iv_j[Q^i_l(w), Q^j_l(w)]^{\Delta_n}_{(\alpha,\beta]} \qquad \text{(by (3.8))}\\
&\quad=\lim_{n\to\infty}[vQ_l(w), vQ_l(w)]^{\Delta_n}_{(\alpha,\beta]}=0 \qquad \text{(by the assumption).}
\end{aligned}$$

It means that

$$vQ_{lk}(\lambda, w)=0, \qquad a\leqslant\lambda\leqslant b, \quad k=1, ..., r. \tag{3.9}$$

Moreover, for any rational numbers α and β such that $a\leqslant\alpha<\beta\leqslant b$,

$$\begin{aligned}
&v\left(Q_l(\beta, w)-Q_l(\alpha, w)-\int_\alpha^\beta Q_{l0}(\lambda, w)\,d\lambda\right)\\
&\quad=\lim_{n\to\infty}\sum_{k=1}^{r} v(I^{\Delta_n}_{k,(\alpha,\beta]}(Q_{lk}))(w) \qquad \text{(by (3.7))}\\
&\quad=\lim_{n\to\infty}\sum_{k=1}^{r}(I^{\Delta_n}_{k,(\alpha,\beta]}(vQ_{lk}))(w)=0 \qquad \text{(by (3.9))}
\end{aligned}$$

and since $vQ_l(\alpha, w)=vQ_l(\beta, w)=0$, we have

$$vQ_{l0}(\lambda, w)=0, \qquad a\leqslant\lambda\leqslant b.$$

Repeating this argument, we have the assertion of the lemma. ∎

Proposition 3.6. *Under the assumptions of Theorem* 3.1, *it holds that if* $w \in \mathbb{W} - N$, *then*

(i) $F_t^1(w), ..., F_t^d(w)$ *are linearly independent for all* $t > 0$,

(ii) $\operatorname{span}(F_t^i(w) : 1 \leqslant i \leqslant d,\ t \leqslant a) \cap \operatorname{span}(F_b^i(w) : 1 \leqslant i \leqslant d) = \{0\}$ *if* $0 \leqslant a < b \leqslant 1$.

Proof. Suppose that $v = (v_1, ..., v_d) \in \mathbb{R}^d$, $0 \leqslant a < b \leqslant 1$, *and* $vF_b(w) = \sum_{i=1}^d v_i F_b^i(w) \in \operatorname{span}(F_t^i(w) : 1 \leqslant i \leqslant d,\ 0 \leqslant t \leqslant a)$. It suffices to show that $v = 0$. Since $\dot{F}_t(\lambda) = 0$ if $t < \lambda$, we have $v\dot{F}_b(\lambda, w) = 0$ if $a \leqslant \lambda \leqslant b$. Therefore

$$v(\dot{F}_{l,b}(w))(\lambda) = vY_b(w)Q_l(\lambda, w) = 0, \qquad a \leqslant \lambda \leqslant b, \quad l = 1, ..., r.$$

and by Lemma 3.5, we have

$$vY_b(w)Q_{lk_1 \cdots k_m}(\lambda, w) = 0, \qquad a \leqslant \lambda \leqslant b, \quad 1 \leqslant l \leqslant r,$$
$$0 \leqslant k_1, ..., k_m \leqslant r.$$

Let us choose $\lambda \in [a, b]$ such that

$$\operatorname{rank}\{\sigma_{k_0 \cdots k_m}(\lambda, x, w) : m \geqslant 0, k_0 \neq 0\} = d \qquad \text{for all} \quad x \in \mathbb{R}^d$$

according to the assumption. Then we have

$$vY_b(w)\,Y_\lambda^{-1}(w)\sigma_{k_0 \cdots k_m}(\lambda, X_\lambda(w), w) = 0 \qquad \text{if} \quad k_0 \neq 0.$$

So, we have $vY_b(w)\,Y_\lambda^{-1}(w) = 0$ and $v = 0$ follows. ∎

Now we shall prove Theorems 3.1 and 3.2.

Proof of Theorem 3.1. Let $\mathbb{H}_0(w)$ be the subspace of $\mathbb{H}$ spanned by the $DX_0^i(w)$'s and $(DB^i(t, x))(w)$'s, which has finite dimension by the assumption. If $t > 0$ and $\operatorname{span}(F_t^i(w), 1 \leqslant i \leqslant d) \cap \mathbb{H}_0(w) = \{0\}$, then $DX_t^1(w), ..., DX_t^d(w)$ must be linearly independent since $F_t^1(w), ..., F_t^d(w)$ are linearly independent (Proposition 3.6(i)) and $DX_t^i(w) - F_t^i(w)$ belongs to $\mathbb{H}_0(w)$ (Proposition 3.4). Therefore if $0 < t_1 < \cdots < t_n \leqslant 1$ and $DX_{t_m}^1(w), ..., DX_{t_m}^d(w)$ are linearly dependent for $m = 1, ..., n$, then there exists a nonzero element $h_m \in \operatorname{span}(F_{t_m}^i(w),\ 1 \leqslant i \leqslant d) \cap \mathbb{H}_0(w)$ for each $m = 1, ..., n$. But $h_1, ..., h_n$ are linearly independent by Proposition 3.6(ii) so $n \leqslant \dim \mathbb{H}_0(w)$. It follows that the number of t's such that $DX_t^1(w), ..., DX_t^d(w)$ are linearly dependent is at most $\dim \mathbb{H}_0(w)$ for each $w \in \mathbb{W}$. We put

$$A_t = \{w \in \mathbb{W} : DX_t^1(w), ..., DX_t^d(w) \text{ are linearly dependent}\}.$$

Then

$$\int_0^1 P(A_t)\,dt = \int_{\mathbb{W}} \left(\int_0^1 \chi_{A_t}(w)\,dt \right) P(dw) = 0$$

so $P(A_t) = 0$ for a.e. $t \in [0, 1]$. Then the first assertion of the theorem follows by Lemma 2.6.

Moreover if $\sup_{w \in \mathbb{W}} \dim \mathbb{H}_0(w) = n < \infty$, $A_{t_1} \cap \cdots \cap A_{t_{n+1}} = \varnothing$ for arbitrary distinct $t_1, \ldots, t_{n+1} \in [0, 1]$ so the number of t's such that $P(A_t) > 0$ is at most countable (Appendix 2). It implies the second assertion of the theorem ∎

Proof of Theorem 3.2. By the assumption, there exists a single stochastic differential equation of type (3.1) which has a solution $\{\xi_t = (\xi_t^1, \ldots, \xi_t^e)\}_{0 \leqslant t \leqslant 1}$ and $t_1, \ldots, t_e \in (0, 1]$ such that

$$D_l X_0^i(w),\ (D_l B^i(t, x))(w) \in \operatorname{span}\{D_k \xi_{t_n}^n(w) : 1 \leqslant n \leqslant e,\ 1 \leqslant k \leqslant r\}$$

for all l, w, i, t, x. Let $G_{l,t}^i$ be defined by (3.6). Then $G_{l,t}^i(w)$ belongs to $\operatorname{span}\{D_k \xi_{t_n}^n(w) : 1 \leqslant n \leqslant e,\ 1 \leqslant k \leqslant r\}$ so it is represented as

$$G_{l,t}^i = \sum_{k,n} v_{l,k,n}^i D_k \xi_{t_n}^n, \tag{3.10}$$

where the $v_{l,k,n}^i$'s are functions on $\mathbb{W}$. Let an $\mathbb{R}^e \otimes \mathbb{R}^e$-valued stochastic process $\{U_t\}_{0 \leqslant t \leqslant 1}$ be the solution to the stochastic differential equation

$$U_t = I + \sum_{k=1}^r \int_0^t \partial V_k(s, \xi_s) U_s \circ dw^k(s) + \int_0^t \partial V_0(s, \xi_s) U_s\,ds$$

and set $R_{k_0 \cdots k_m}(t) = U_t^{-1} V_{k_0 \cdots k_m}(t, \xi_t)$. Then we have

$$(\dot{D}_l \xi_t)(\lambda) = U_t R_l(\lambda) \chi_{[0,t]}(\lambda), \qquad 0 \leqslant \lambda \leqslant 1$$

which is the Markovian case in Proposition 3.4. We shall fix an arbitrary $t \in (0, 1]$ and choose an ε such that $0 < \varepsilon < \min(t, t_1, \ldots, t_e)$. Then, by Proposition 3.4 and (3.10), we have

$$(\dot{D}_l X_t)(\lambda) = Y_t Q_l(\lambda) + \sum_{k,n} v_{l,k,n} U_{t_n} R_k(\lambda), \qquad 0 \leqslant \lambda \leqslant \varepsilon.$$

By a similar argument to the proof of Lemma 3.5, choosing a good version for the $Q_{k_0 \cdots k_m}$'s and $R_{k_0 \cdots k_m}$'s, we see that

$$w \in \mathbb{W}, \quad v \in \mathbb{R}^d, \quad \tilde{v}_k \in \mathbb{R}^e, \quad 1 \leqslant k \leqslant r, \qquad \text{and}$$

$$v Q_l(\lambda, w) + \sum_k \tilde{v}_k R_k(\lambda, w) = 0, \qquad 0 \leqslant \lambda \leqslant \varepsilon$$

imply

$$vQ_{lk_1\cdots k_m}(\lambda, w) + \sum_k \tilde{v}_k R_{kk_1\cdots k_m}(\lambda, w) = 0, \qquad 0 \leqslant \lambda \leqslant \varepsilon$$

for $m \geqslant 1$ and $0 \leqslant k_1, ..., k_m \leqslant r$.

Suppose that $DX_t^1(w), ..., DX_t^d(w)$ are linearly dependent so that there exists a nonzero vector $v = (v_1, ..., v_d) \in \mathbb{R}^d$ such that $v \cdot DX_t(w) = \sum_i v_i DX_t^i(w) = 0$. Then since $v(\dot{D}_l X_t(w))(\lambda) = 0$ for $0 \leqslant \lambda \leqslant \varepsilon$ and $l = 1, ..., r$, we have

$$vY_t(w)Q_l(\lambda, w) + v\sum_{k,n} v_{l,k,n}(w)U_{t_n}(w)R_k(\lambda, w) = 0, \qquad 0 \leqslant \lambda \leqslant \varepsilon$$

for $l = 1, ..., r$. Therefore

$$vY_t(w)Q_{lk_1\cdots k_m}(\lambda, w) + v\sum_{k,n} v_{l,k,n}(w)U_{t_n}(w)R_{kk_1\cdots k_m}(\lambda, w).$$

Put $\lambda = 0$ and note that $R_{k_0\cdots k_m}(0) = V_{k_0\cdots k_m}(0, \xi) = 0$ whenever $m \geqslant q$. Then we have

$$vY_t(w)\sigma_{k_0\cdots k_m}(0, X_0(w), w) = 0, \qquad m \geqslant q, \quad 1 \leqslant k_0 \leqslant r,$$
$$0 \leqslant k_1, ..., k_m \leqslant r.$$

Since the vector $vY_t(w)$ is nonzero, we have

$$\operatorname{rank}\{\sigma_{k_0\cdots k_m}(0, X_0(w), w) : m \geqslant q, k_0 \neq 0\} < d.$$

But the totality of such w is a null set by the assumption. Therefore $DX_t^1, ..., DX_t^d$ are linearly independent a.s. and noting Lemma 2.6, the assertion of Theorem 3.2 follows. ∎

Appendix 1

We show a property concerning the paths of the stochastic integral as a process, which is used to show the important properties (3.7) and (3.8).

Proposition. *Let $\{u(t)\}_{0 \leqslant t \leqslant 1}$ and $\{v(t)\}_{0 \leqslant t \leqslant 1}$ be continuous processes which belong to $\tilde{\mathbb{L}}^{1,2}_{loc}(\mathbb{R})$. We put $\xi_t = \int_0^t u(s) \circ dw^k(s)$ and $\eta_t = \int_0^t v(s) \circ dw^l(s)$ for some $k, l = 1, ..., r$. Let $0 \leqslant a < b \leqslant 1$. Then,*

$$[\xi, \eta]^{\Delta}_{(a,b]} \to \delta_{kl} \int_a^b u(t)v(t)\, dt \qquad \text{in probability}$$

as $|\Delta| \to 0$ *where we set*

$$[\xi, \eta]^{\Delta}_{(a,b]} = \sum_{m=1}^{n} (\xi_{t_m \wedge b} - \xi_{t_{m-1} \vee a})(\eta_{t_m \wedge b} - \eta_{t_{m-1} \vee a})$$

for a partition $\Delta : 0 = t_0 < t_1 < \cdots < t_n = 1$ *of* $[0, 1]$.

Proof. We assume that $u = v$ and $k = l$ since the other cases can be handled in the same way. We write $w^k = w$ and $D_k = D$ for simplicity. We have the representation

$$\int_0^t u(s) \circ dw(s) = \int_0^t u(s)\, dw(s) + a(t),$$

where $\int_0^t u(s)\, dw(s)$ denotes the Skorokhod integral and $\{a(t)\}_{0 \leq t \leq 1}$ is a process which has absolutely continuous paths (cf. [4]). For any partitions $\Delta : a = t_0 < t_1 < \cdots < t_n = b$ of $[a, b]$,

$$\begin{aligned}
\int_{t_{m-1}}^{t_m} u(s) \circ dw(s) &= \int_{t_{m-1}}^{t_m} u(s)\, dw(s) + (a(t_m) - a(t_{m-1})) \\
&= \int_{t_{m-1}}^{t_m} \left(\frac{1}{t_m - t_{m-1}} \int_{t_{m-1}}^{t_m} u(t)\, dt \right) dw(s) \\
&\quad + \int_{t_{m-1}}^{t_m} \left(u(s) - \frac{1}{t_m - t_{m-1}} \int_{t_{m-1}}^{t_m} u(t)\, dt \right) dw(s) \\
&\quad + (a(t_m) - a(t_{m-1})).
\end{aligned}$$

By applying the formula $\int_a^b F\, dw(t) = F(w(b) - w(a)) - \int_a^b (\dot{D}F)(\lambda)\, d\lambda$ concerning $F \in \mathbb{D}^{1,2}$ (cf. [4]), we have

$$\begin{aligned}
&= \left(\frac{1}{t_m - t_{m-1}} \int_{t_{m-1}}^{t_m} u(t)\, dt \right) (w(t_m) - w(t_{m-1})) \\
&\quad - \frac{1}{t_m - t_{m-1}} \int_{t_{m-1}}^{t_m} \int_{t_{m-1}}^{t_m} (\dot{D}u(t))(\lambda)\, dt\, d\lambda \\
&\quad + \int_{t_{m-1}}^{t_m} \left(u(s) - \frac{1}{t_m - t_{m-1}} \int_{t_{m-1}}^{t_m} u(t)\, dt \right) dw(s) \\
&\quad + (a(t_m) - a(t_{m-1})) \\
&\equiv T^{\Delta}_{1,m} + T^{\Delta}_{2,m} + T^{\Delta}_{3,m} + T^{\Delta}_{4,m}.
\end{aligned}$$

Let $|\Delta| \to 0$. We shall show in five steps that $[\xi, \xi]^{\Delta}_{(a,b]} \to \int_a^b u(t)^2\, dt$ in probability. We need the continuity of the paths of $\{u(t)\}_{0 \leq t \leq 1}$ only in the fourth step.

(I) It follows from Schwarz' inequality that

$$\sum_m |T_{2,m}^{\Delta}|^2 \leqslant \sum_{m=1}^{n} \int_{t_{m-1}}^{t_m} \int_{t_{m-1}}^{t_m} |(\dot{D}u(t))(\lambda)|^2 \, dt \, d\lambda$$
$$\to 0$$

since $\int_0^1 \|Du(t)\|_H^2 \, dt < \infty$ a.s. (see Definition 2.1).

(II) Similarly, we have $\sum_m |T_{4,m}^{\Delta}|^2 \to 0$ a.s.

(III) We may assume that $\{u(t)\}_{0 \leqslant t \leqslant 1}$ belongs to $\tilde{\mathbb{L}}^{1,2}$. We have

$$T_{3,m}^{\Delta} = \int_{t_{m-1}}^{t_m} \{u(t) - u^{\Delta}(t)\} \, dw(t),$$

where we define the stochastic process $\{u^{\Delta}(t)\}_{a \leqslant t \leqslant b}$ by

$$u^{\Delta}(t) = \frac{1}{t_m - t_{m-1}} \int_{t_{m-1}}^{t_m} u(t) \, dt, \qquad \text{if} \quad t_{m-1} < t \leqslant t_m.$$

Then,

$$\begin{aligned}
\mathbb{E}\left[\sum_m |T_{3,m}^{\Delta}|^2\right] &= \sum_m \mathbb{E}\left[\left|\int_{t_{m-1}}^{t_m} \{u(t) - u^{\Delta}(t)\} \, dw(t)\right|^2\right] \\
&\leqslant \sum_m \int_{t_{m-1}}^{t_m} \|u(t) - u^{\Delta}(t)\|_{\mathbb{D}^{1,2}}^2 \, dt \qquad \text{(see [4])} \\
&= \int_a^b \|u(t) - u^{\Delta}(t)\|_{\mathbb{D}^{1,2}}^2 \, dt \to 0.
\end{aligned}$$

(IV) We shall show that $\sum_m |T_{1,m}^{\Delta}|^2 \to \int_a^b u(t)^2 \, dt$ in probability. Let $\{\Delta^{(j)}\}_{\alpha}^{\infty} = 1$ be an arbitrary sequence of partitions of $[a, b]$ such that $|\Delta^{(j)}| \to 0$. Then, $P(\sup_{a \leqslant t \leqslant b} |t - a - [w]_{(a,t]}^{\Delta^{(j)}}| \geqslant \varepsilon) \to 0$ for any $\varepsilon > 0$ where $[w]_{(a,t]}^{\Delta} = \sum_{m=1}^{n} \{w(t_m \wedge b) - w(t_{m-1} \vee a)\}^2$ for $\Delta : a = t_0 < \cdots < t_n = b$ so that there exists a subsequence $\{\tilde{\Delta}^{(j)}\}_{j=1}^{\infty}$ such that

$$\sup_{a \leqslant t \leqslant b} |t - a - [w]_t^j| \to 0 \qquad \text{a.s.},$$

where $[w]_t^j = [w]_{(a,t]}^{\tilde{\Delta}^{(j)}}$. It means that the Stieltjes measure $d[w]_t^j$ on $(a,b]$ converges to the Lebesgue measure weakly. Then, putting $u^{\tilde{\Delta}^{(j)}} = u_j$, we have

$$\left|\sum_m |T_{1,m}^{\Delta(j)}|^2 - \int_a^b u(t)^2\,dt\right|$$

$$= \left|\int_a^b u_j(t)^2\,d[w]_t^j - \int_a^b u(t)^2\,dt\right|$$

$$\leqslant \int_a^b |u_j(t)^2 - u(t)^2|\,|d[w]_t^j| + \left|\int_a^b u(t)^2\,d[w]_t^j - \int_a^b u(t)^2\,dt\right|$$

$$\leqslant \Big(\sup_{a\leqslant t\leqslant b} |u_j(t)^2 - u(t)^2|\Big)[w]_b^j$$

$$+ \left|\int_a^b u(t)^2\,d[w]_t^j - \int_a^b u(t)^2\,dt\right|$$

$$\to 0 \qquad \text{as} \quad j\to\infty$$

since u is continuous. It means that $\sum_m |T_{1,m}^{\Delta}|^2 \to \int_a^b u(t)^2\,dt$ in probability.

(V) It follows from (I)–(IV) that

$$[\xi,\xi]_{(a,b]}^{\Delta} = \sum_m |\xi(t_m) - \xi(t_{m-1})|^2 = \sum_m (T_{1,m}^{\Delta} + \cdots + T_{4,m}^{\Delta})^2$$

$$= \sum_{i,j=1}^{4} \sum_m T_{i,m}^{\Delta} T_{j,m}^{\Delta} \to \int_a^b u(t)^2\,dt \qquad \text{in probability,}$$

where we can see that the cross terms vanish since $|\sum_m T_{i,m}^{\Delta} T_{j,m}^{\Delta}|^2 \leqslant (\sum_m |T_{i,m}^{\Delta}|^2)(\sum_m |T_{j,m}^{\Delta}|^2)$ by Schwarz' inequality. ∎

APPENDIX 2.

We shall prove the following set-theoretical property which we used in the proof of Theorem 3.1.

PROPOSITION. *Let $\mathscr{A}$ be a class of measurable subsets of $\mathbb{W}$. We assume that there exists a natural number n such that $P(A_1 \cap \cdots \cap A_n) = 0$ for any distinct elements $A_1, ..., A_n$ of $\mathscr{A}$. Then the number of elements A of $\mathscr{A}$ such that $P(A) > 0$ is at most countable.*

Proof. It suffices to show that $\sum_{j=1}^{q} P(A_j) \leqslant n-1$ for any finitely many distinct elements $A_1, ..., A_q$ of $\mathscr{A}$. For given $A_1, ..., A_q$, we set

$$\tilde{A}_j^1 = A_j$$

$$\tilde{A}_j^m = \bigcup_{j_1 < \cdots < j_{m-1} < j} (A_{j_1} \cap \cdots \cap A_{j_{m-1}} \cap A_j) \qquad \text{if} \quad 2 \leqslant m \leqslant j$$

$$\tilde{A}_j^m = \varnothing \qquad \text{if} \quad m > j.$$

$$A_j^m = \tilde{A}_j^m - \tilde{A}_j^{m+1}, \qquad 1 \leqslant j \leqslant q, \quad m = 1, 2, \ldots.$$

Then, $A_j^1, A_j^2, \ldots$ are disjoint, $A_j = \bigcup_m A_j^m$, and $P(A_j^n) = P(A_j^{n+1}) = \cdots = 0$ by the assumption. Therefore,

$$P(A_j) = \sum_{m=1}^{n-1} P(A_j^m), \qquad j = 1, \ldots, q$$

holds. On the other hand, since $\tilde{A}_j^m \cap A_{j+1} \subset \tilde{A}_{j+1}^{m+1}$ and $\tilde{A}_{j+1}^{m+1} \cap A_{j+1}^m = \varnothing$, we have $A_i^m \cap A_{j+1}^m \subset \tilde{A}_j^m \cap A_{j+1}^m = \varnothing$ so that

$$\sum_{j=1}^{q} P(A_j^m) = P\left(\bigcup_j A_j^m\right) \leqslant 1, \qquad m = 1, 2, \ldots.$$

Therefore $\sum_{j=1}^q P(A_j) = \sum_{j=1}^q \sum_{m=1}^{n-1} P(A_j^m) \leqslant n - 1.$ ∎

References

1. J. M. Bismut, Martingales, the Malliavin calculus and hypoellipticity under general Hörmander's conditions, *Z. Wahrsch. Verw. Gebiete* **56** (1981), 469–505.
2. S. Kusuoka and D. Stroock, Applications of the Malliavin calculus, II, *J. Fac. Sci. Univ. Tokyo Sect. IA Math.* **34** (1985), 1–76.
3. P. Malliavin, Stochastic calculus of variations and hypoelliptic operators, *in* "Proceedings, International Conf. on S.D.E. at Kyoto, 1976," pp. 195–263, Wiley/Kinokuniya, New York/Tokyo, 1978.
4. D. Nualart and E. Pardoux, Stochastic calculus with anticipating integrands, *Probab. Theory Related Fields* **78** (1988), 35–581.
5. D. Ocone and E. Pardoux, Generalized Ito–Ventzell formula. Application to a class of anticipating stochastic differential equations, *Ann. Inst. H. Poincaré* **25** (1989), 39–71.
6. I. Shigekawa, Derivatives of Wiener functionals and absolute continuity of induced measures, *J. Math. Kyoto Univ.* **25** (1980), 263–289.
7. S. Watanabe, "Lectures on Stochastic Differential Equations and Malliavin's Calculus," Tata Institute of Foundamental Research, Springer, Berlin/Heidelberg/New York, 1984.